Real-time motion capture and animation generation in virtual character using Kinect for American Sign Language (ASL)
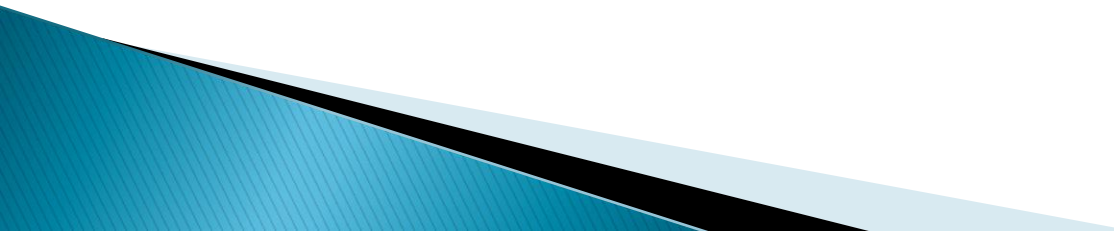
Affective Intelligent Agents (CAP 5627)
  -Ahamed, Md Shakil
  -Professor: Christine Lisetti, PhD

FIU
**Computing &
Information
Sciences**

Outline

- ❑ Introduction
  - • Motivation
  - • Problem Statement
- ❑ Related Research
- ❑ Methodologies

Introduction

- Smooth gesture animation is very important for getting meaningful ASL translation in virtual character
- Kinect device controlled motion capture can play role for getting smooth animation

virtual character

Kinect Device

Introduction- ASL

| American Sign Language statistics	
Region	North America, Canada, West & Central Africa
Native Speaker	250,000-500,000 in the US (1972)
Language Family	French sign language
Origin	19th Century, Connecticut
Source	[Mitchell, et al, 2006].

Introduction- Kinect Motion Capture

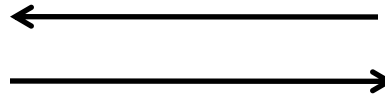# Introduction- Smartbody character (ICT 2004)

Introduction- Virtual character in ASL

virtual character/
Translator

Doctor

Signer Patient
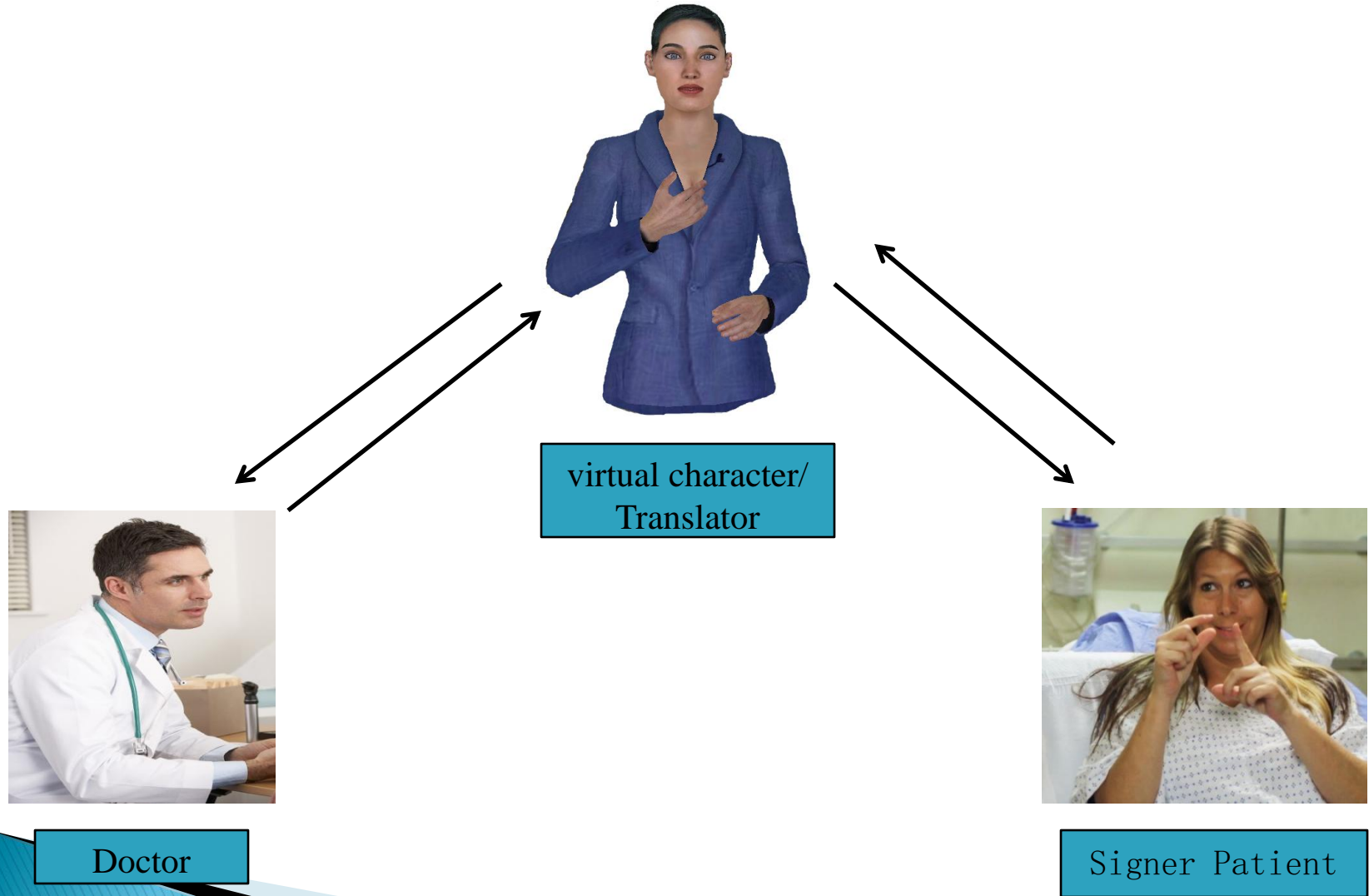
Introduction- Virtual character in ASL

virtual character/
Teacher

Signer Student

Introduction-virtual character in daily life

virtual character

Daily Life

Gaming · Go to bed · Fall Asleep · Sleep · Sleep · Sleep · Sleep · Wake up · Actually wake up · Get ready · Go to college · Class · Class · Class · Lunch · Class · Class · Go home · Internet · Gaming · Dinner · Gaming · Gaming
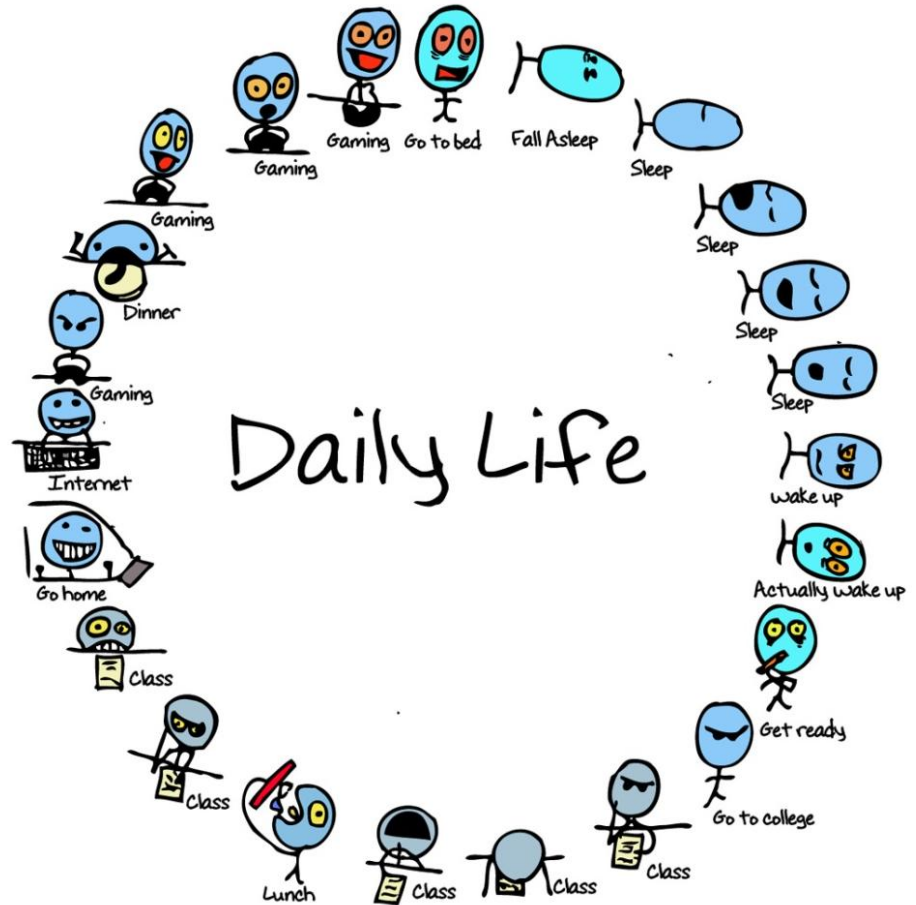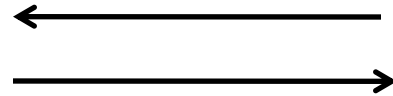
Motivation

- ❑ Develop real-time gesture animation controller
- ❑ Develop direction of movement codes for ASL Translator for Affective Computing Lab
- ❑ Develop platform for motion corpus data extraction
- ❑ Smartbody opensource animation platform

Story Teller

Problem Statement

- Definition: Generate smooth direction of movement animation using motion capture data
- Why?
  - Hard to get many gesture with limited time
  - Generating smooth timebased animation is time consuming
- How?
  - Using Kinect motion capture

Related research

- ASL MoCap capture research by CUNY
  - Created motion capture corpus data
  - First released in 2012.
  - Equipment:
    - 3 HD video cameras, cyber gloves, head mounted eyetracker,
    - An intersense IS-900 acoustic tracker (for location and orientation of the signer's head)
  - Issues
    - Devices are expensive
   - Managing the equipment's are difficult

Pengfei, et al 2012

Related research

- Chinese Sign to Text translation using Kinect
  - 3D trajectory matching algorithm based on a set of gallery images
  - Find the best matching trajectory from the gallery
  - Used Kinect device!!
- Issues
  - Used manual (not automatic) animation for the words
  - They mainly focuses on sign recognition part pattern of signs

X Chai, et al, 2013

Related research

- ❑ Animation & Comprehensibility
  - ◆ Developed signle gloss based animation system
  - ◆ Focussed on animation understandibility
    - ◆ focused on quality in time of generation
    - ◆ Evaluation
      - ◆ Subjective-video remake quality testing & verification
      - ◆ Objective - # of repeatiion asked by subjects
  - ◆ Used tools:
    - ◆ OpenMARY-TTS system
    - ◆ ANVIL annotation tool- for video segmentation
    - ◆ Behavior builder tool – support single and sequence of glesses animation
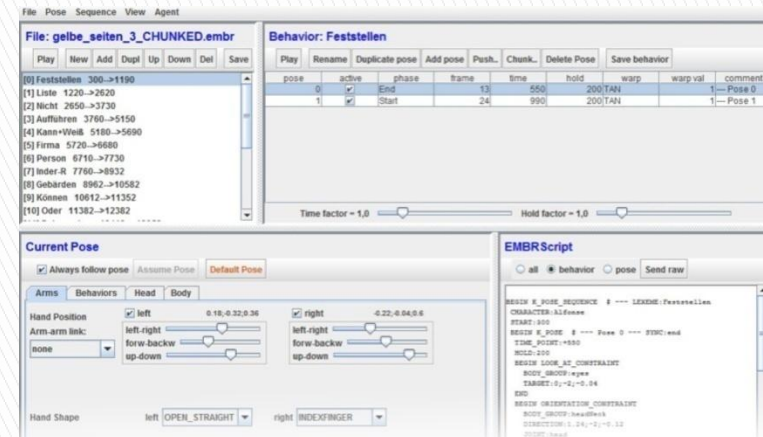
- ❑ Issues:
  - ◆ Comprehensibility 58.6%
  - ◆ Created #95 glosses
  - ◆ Need manual intervention for sign correction
  - ◆ Ignore grammatical flections, Misses each gloss related information

Kipp, et al 2011

Related research

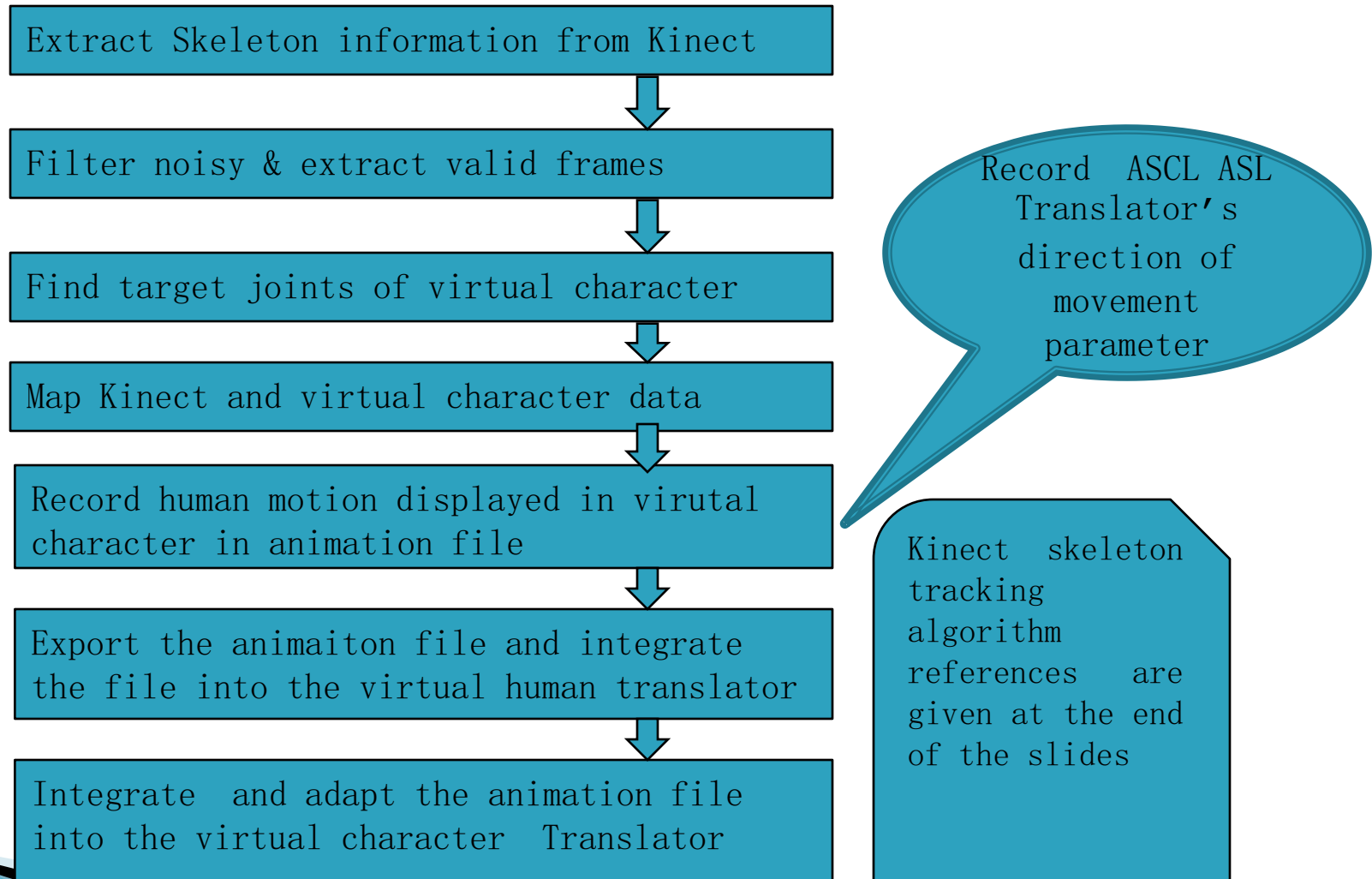- Survey of English to ASL animation
  - M. huernauth studied ASL tranlators upto 2003
  - Studied 4 projects and showed comparative analysis:
    - Vicicast
    - Zardos
    - Team project
    - ASL workbench
  - Team system performed better animation;
  - All used manual animation (timeline based) techniques
  - Their main foucs were on English to ASL conversion, grammar generation, maintain semantic ordering

Related research

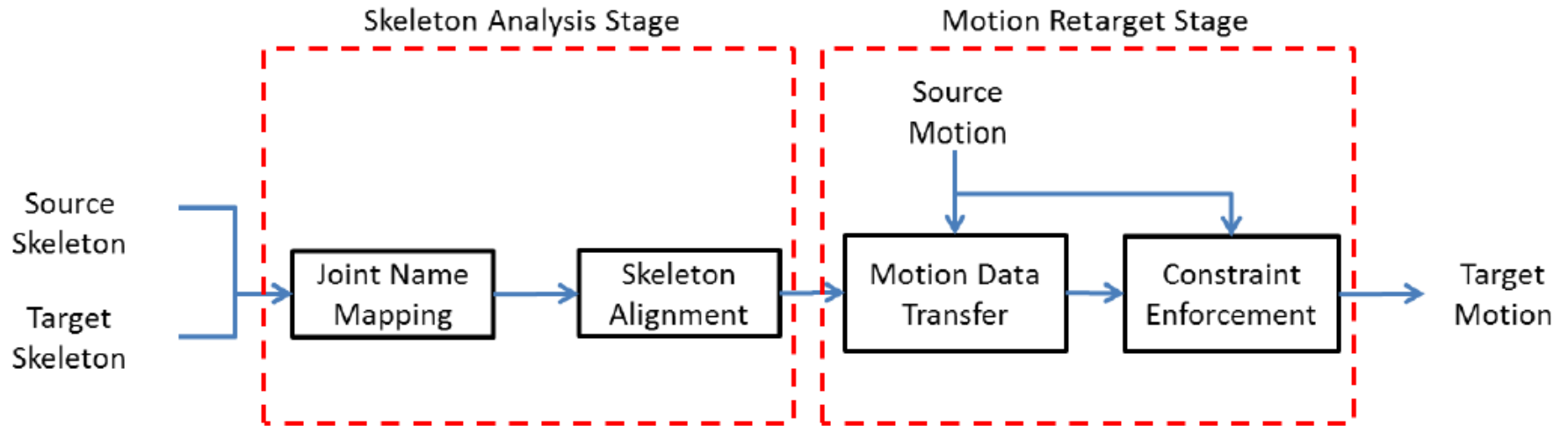- Analysis of ASL motion capture data towards Identification of Verb Type
    - Built for sign recognition application
    - Captured the motion data and classify the data into 2 major categories telic and atelic based on captured motions velocity and acceleration
    - Telic words: send, happen, hit (has end point)
    - Atelic words: play, read, run
- Equipment : Many Mocap devices requires!!
    - Gypsy 3.0 wired motion capture suit
    - Pair of 18-sensor Cybergloves.
    - Six motion capture cameras.
    - Motion blender software
   - Elan software for data annotation

Malaia, et al, 2008

Methodology

Extract Skeleton information from Kinect

↓

Filter noisy & extract valid frames

↓

Find target joints of virtual character

↓

Map Kinect and virtual character data

↓

Record human motion displayed in virutal character in animation file

↓

Export the animaiton file and integrate the file into the virtual human translator

↓

Integrate and adapt the animation file into the virtual character Translator

Record ASCL ASL Translator's direction of movement parameter

Kinect skeleton tracking algorithm references are given at the end of the slides

Skeleton & Motion retargetting

From Smartbody Manual

Data mapping issues

- ❑ Kinect Dude character vs target virtual character
  - ➢ Joint names between two character may be different
  - ➢ Different initial pose
  - ➢ Different local rotation frame e.g. 10 unit change may differ between character
  - ➢ Recorded file format issues as character might be in different platform
  - ➢ Kinect data filtering before mapping
  - ➢ Different proportion and scale between two characters
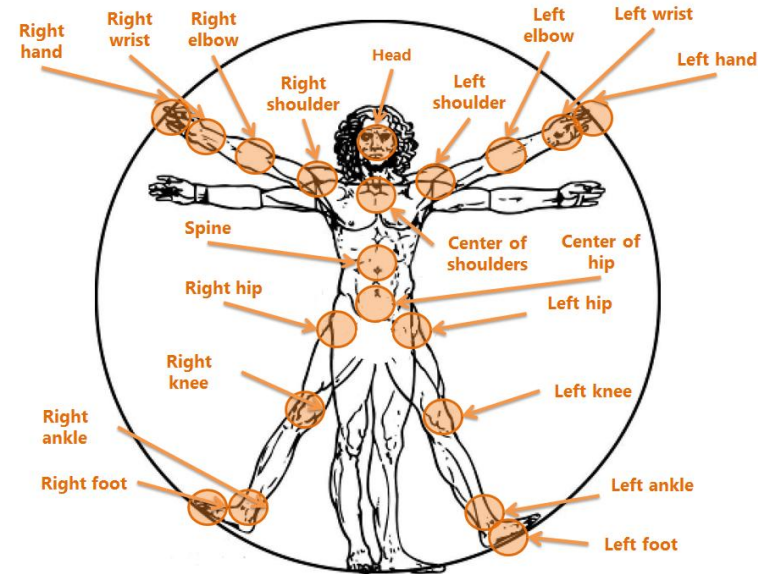
Current limitations

- ❑ Target character: Smartbody
  - Have project build issues e.g. Time
  - Project size are not optimized
  - For character animation they depended on Maya (not open source)
  - Used many 3rd party softwares some of them are not open source
  - ****SB Forum discueses FbxToSbConverter.exe tool which onverts fbx file to .skm file format

Next plan of work

- Motion record using Smartbody character
  - Solve build issues of Smartbody (time, size)
  - Figure out how to record valid captured frame
  - Record the gestures data for ASCL ASL project
  - Make the performance evaluation how the animation data works

References

- X Chai, et al, "Sign Lanugage Recognition and Translation with kinect", IEEE Conf. on AFGR, 2013.
- Pengfei Lu, et al, "CUNY American Sign Language Motion-Capture Corpus: First Release." Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon, The 8th International Conference on Language Resources and Evaluation, Istanbul, Turkey, 2012.
- Evguenia Malaia, et al, "Analysis of ASL motion capture data towards identification of verb type",Proceeding STEP '08 Proceedings of the 2008 Conference on Semantics in Text Processing Pages 155-164, Association for Computational Linguistics Stroudsburg, PA, USA, 2008
- Matt Huenerfauth. 2003. A Survey and Critique of American Sign Language Natural Language Generation and Machine Translation Systems. Technical Report MS-CIS-03-32, Computer and Information Science, University of Pennsylvania.
- K Mulrooney, "American Sign Language DeMystified", Book.
- Pengfei Lua, Matt Huenerfauth, "Collecting and evaluating the CUNY ASLcorpus for research on American Sign Language animation",ScienceDirect, Computer Speech and Language 28 (2014) 812–831.
- Shotton, et al,"Real-Time Human Pose Recognition in Parts from Single Depth Images", Communications of the ACM, Volume 56 Issue 1, Pages 116-124 , USA, 2013
- Michael Kipp, et al, "Sign Language Avatars: Animation and Comprehensibility", Intelligent Virtual Agents Lecture Notes in Computer Science Volume 6895, 2011, pp 113-126, 2011

Q & A